

Universiti Teknologi MARA

**Sentiment Analysis of Airline Reviews
Using Naïve Bayes Algorithm**

Ahmad Firdaus Bin Maliki

**Thesis submitted in fulfilment of the requirements
for Bachelor of Computer Science (Hons.) Faculty
of Computer and Mathematical Sciences**

FEBRUARY 2021

ACKNOWLEDGMENT

Alhamdulillah, praises, and thanks to Allah because of His Almighty and His utmost blessings, I was able to finish this research within the time duration given. Firstly, my special thanks to my supervisor, Sir Ahmad Nazmi Bin Fadzal for always giving me his full support and supervised me until this project managed to be accomplished on time.

Besides, I would like to give a special thanks to my CSP650 lecturer, Madam Norlina Binti Mohd Sabri for teaching me through this semester and sharing all the knowledge to ensure that I complete this research on time.

Special appreciation also goes to my beloved parents for all the endless prayers, giving me full support, and always believing in me in completing this research successfully and excellently. Last but not least, I would like to give my gratitude to my dearest friend who has helped and gave me their support throughout this semester.

ABSTRACT

Reviews and customer's feedback play a vital role for companies to monitor their brand, improve customer service or gain customer's trust. Even though technologies have evolved over the years, some companies such as airline companies still use traditional methods to collect reviews by using surveys which may lead to human error. This includes bias and random response. Furthermore, with the advancement of technology, people tend to express their satisfaction or dissatisfaction through social media such as Facebook or Twitter which is becoming more popular even among the older generations. Therefore, this project aims to help the company to classify and visualize the sentiments of the reviews by collecting the reviews from Twitter or uploading a file containing reviews to the system. The sentiment of the reviews will be classified into two categories which are positive and negative. The overview for the methodology is divided into three phases which are preliminary phase, design and implementation phase and evaluation phase. Several classifier models have been built by using Naïve Bayes algorithm during the design and implementation phase where the model that has the highest accuracy has been chosen for this project. On top of that, functionality testing has been done on the system to ensure that there are no problems especially in the visualization component of the system. Some recommendations have also been given to improve the project such as training the model with other languages such as Malay and analyse reviews from other social media such as Facebook. In conclusion, this project may help to ease the burden of airline companies as they no longer need to analyse the reviews one by one.

TABLE OF CONTENTS

CONTENT	PAGE
SUPERVISOR APPROVAL	ii
STUDENT DECLARATION	iii
ACKNOWLEDGMENT	iv
ABSTRACT	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	viii
LIST OF TABLES	x
CHAPTER ONE: INTRODUCTION	
1.1 Background of Study	1
1.2 Problem Statement	2
1.3 Project Objectives	3
1.4 Project Scope	3
1.5 Project Significance	3
1.6 Overview of Research Framework	4
1.7 Summary	5
CHAPTER TWO: LITERATURE REVIEW	
2.1 Sentiment Analysis	6
2.2 Airline Review	8
2.3 Naïve Bayes Algorithm	10
2.3.1 Naïve Bayes Classifier	10
2.3.2 Bayes Theorem	11
2.3.3 Bayes Theorem Formula	11
2.3.4 Types of Naïve Bayes Classifier	12
2.4 Implementation of Naïve Bayes Algorithm in Various Problems	13
2.5 Similar Works	16
2.6 The Implication of Literature Review	18
2.7 Conclusion	19

CHAPTER THREE: METHODOLOGY

3.1	Overview of Research Methodology Framework	20
3.1.1	Detailed of Research Methodology Framework	21
3.2	Preliminary Study	24
3.2.1	Literature Study	24
3.3	Data Analysis	25
3.3.1	Data Collection	25
3.3.2	Data Pre-processing	27
3.4	Naïve Bayes Algorithm	28
3.5	Design Phase	29
3.5.1	System Architecture	29
3.5.2	Flowchart	32
3.5.3	Design Interface	33
3.5.4	Prototype Implementation	35
3.6	Evaluation Phase	35
3.6.1	System Evaluation	35
3.6.2	Classifier Evaluation	35
3.7	Conclusion	36

CHAPTER FOUR: RESULT AND FINDING

4.1	Conceptual Framework	38
4.2	Program Codes	40
4.2.1	Data Preprocessing	40
4.2.2	Algorithm	43
4.2.3	Classifier Model Creation	44
4.2.4	System	48
4.3	Prototype Interface	55
4.4	Evaluation Results	62
4.4.1	Model Evaluation	62
4.4.2	System Evaluation	66
4.5	Conclusion	69

CHAPTER FIVE: CONCLUSION AND RECOMMENDATION

5.1	Summary of Project	71
5.2	Project Contribution	72
5.3	Project Limitation	72
5.4	Project Recommendation	72
5.5	Conclusion	73

REFERENCES	74
-------------------	----

APPENDICES	78
-------------------	----

APPENDIX A: GANTT CHART	78
--------------------------------	----